Abstract

Systems and processes for more accurate mortgage scoring are described. A proportional hazards model is employed in which not only the occurrence of an event, but also the time to an event such as default of a loan, is considered. In this approach, a hazard rate can be viewed as the chance that an observation will experience an event in the next instant. There are two components to the response, and a binary variable is utilized to indicate whether the event was observed or not, and a time variable. As a result, the number of loans used for modeling is greatly increased, and the time it takes to observe the event, a valuable piece of information in itself, is included in the process. In addition, nonlinear effects are advantageously modeled in a continuous fashion using hat functions to map a series of independent variables. This approach typically yields smaller prediction errors near boundary points.